

РАЗРАБОТКА ИММЕРСИВНОГО ИГРОВОГО ПРОСТРАНСТВА В ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ С ИСПОЛЬЗОВАНИЕМ МУЛЬТИМОДАЛЬНОГО ИНТЕРФЕЙСА

СЕНИЛОВ Михаил Андреевич
доктор технических наук, профессор
ГАЗИЗУЛИН Марат Русланович
студент

ФГБОУ ВО «Ижевский государственный технический университет им. М.Т. Калашникова»
г. Ижевск, Россия

Рассматривается новый подход в создании иммерсивных игровых виртуальных пространств, заключающийся в разработке современного повышающего иммерсивность мультимодального интерфейса, использующего принцип голосового ввода и обработки жестов рук. При создании иммерсивного пространства используются передовые технологии, обуславливающие синергию элементов мультимодального интерфейса, обеспечивающую новый уровень качества программного обеспечения для виртуальной реальности.

Ключевые слова: виртуальная реальность, иммерсивность, мультимодальный интерфейс, обработка жестов рук, распознавание голоса, разработка компьютерных игр.

С развитием технологий развивается и программное обеспечение. Отрасль виртуальной реальности имеет высокий порог вхождения, обусловленный необходимостью освоения и применения пользователем сверхсовременного программного и аппаратного обеспечения, связанного с большими затратами времени и усилий для освоения управления и ориентирования в виртуальном пространстве. При этом в зависимости от разработчика создаются различные по удобству и доступности приложения. Исследования и разработки в данном направлении представляются несомненно актуальными, так как создание виртуального окружения, сравнимого с реальным миром, является одной из самых последних разработок в информационной эре [8, с. 745].

Объектом исследований в данной работе является мультимодальный интерфейс в иммерсивном игровом виртуальном пространстве. Чтобы повысить дружелюбность интерфейса, сделать программное обеспечение для виртуальной реальности более общедоступным и иммерсивным, предлагается пересмотреть средства ввода пользователя. Цель работы заключается в разработке нового подхода к созданию иммерсивных игровых виртуальных пространств. Вместо

управления персонажем клавиатурой, мышью или контроллером предлагается использование голосового ввода, его распознавание, а также отслеживание положения кистей и пальцев рук. Исследования связаны с анализом сферы виртуальной реальности, проводятся с использованием современных VR-библиотек и имеют целью синтез модели мультимодального интерфейса, дружелюбного пользователю. Таким образом, разрабатываемый новый способ взаимодействия пользователей с виртуальной реальностью открывает возможность к повышению дружелюбности и иммерсивности создаваемого программного обеспечения.

По сравнению с работами [2] и [10], в которых элементы интерфейса рассматриваются изолированно друг от друга, в данном исследовании предлагается совмещение распознавания голоса с анализом жестов рук, приводящее к синергетическому эффекту, качественно повышающему дружелюбность мультимодального интерфейса.

Самая важная часть во взаимодействиях в виртуальной реальности – это человек, их выполняющий. Соответственно, так как пользователь – главная фигура, ему нужно уделять особое внимание. Когда виртуальная реальность сделана качественно, взаимодействия

могут быть удобными и приносить удовольствие, в противном случае человек может испытывать фрустрацию, усталость и болезненное самочувствие. Чтобы улучшить пользовательский опыт, нужно понимать человеческое восприятие, учитывать интуитивные взаимодействия [5, с. 431]. Для перехода на качественно новый уровень нужно вернуться к фундаментальным принципам виртуальной реальности и пересмотреть их, добавив новые элементы, например, новое средство ввода, которое не будет принуждать пользователя к физическому взаимодействию, что решается использованием мультимодального интерфейса.

Реализация мультимодального интерфейса включает в себя распознавание речи и отслеживание жестов. Ядром распознавания

речи в выбранной среде разработки Unity является класс `SpeechRecognitionEngine`, использующий `Windows Speech SDK`. Он позволяет использовать анализаторы речи разных видов: распознавание каждого сказанного слова (`OpenDictation`), распознавание заранее сформированного набора слов (`KeywordRecognizer`) или фраз (`PhraseRecognizer`) с заданным уровнем уверенности, от которого зависит погрешность, по принципу ближайшего совпадения [7, с. 857]. Фрагменты кода представлены на языке программирования C#.

Для распознавания речи объявляется анализатор типа `KeywordRecognizer`, настройка и запуск которого происходит на старте программы, инициализация которого представлена на рисунке 1.

```
rec = new KeywordRecognizer(keywords, confidence);
rec.OnPhraseRecognized += Rec_OnRec;
rec.Start();
```

Рисунок 1. Инициализация анализатора речи

После срабатывает событие, код которого изображен на рисунке 2, активирующее

действие, соответствующее тексту распознанной фразы.

```
private void Rec_OnRec(PhraseRecognizedEventArgs args)
{
    word = args.text;
    wordAnalyzer.spellUsed[word] = true;
    wordAnalyzer.keywords[word].Invoke();
}
```

Рисунок 2. Событие, связывающее распознанный текст с действием

Далее реализовано отслеживание рук. Оно позволяет использовать руки как средство ввода для шлемов виртуальной реальности `Oculus Quest`. Выбор рук как средства ввода обеспечивает новое ощущение присутствия, повышает вовлеченность и обеспечивает более естественное взаимодействие с виртуальным миром.

Отслеживание рук анализирует дискрет-

ные позы рук и отслеживает позицию определенных ключевых точек на руке, таких как суставы и кончики пальцев, в реальном времени, пока двигаются руки [1, с. 331]. Когда руки используются как средство ввода, ладонь управляет лазерным курсором-указателем, который ведет себя, как стандартный курсор контроллера. Этот курсор-указатель может использоваться для выделе-

ния, выбора, клика или записи своей событийной логики в приложении. Интегрированные руки могут взаимодействовать с предметами, используя жесты рук: можно указать, ущипнуть, сжать ладонь, и так далее.

Отслеживание и анализ становится возможным благодаря использованию Oculus

SDK. Подключив его в игровом движке Unity, мы объявляем две руки, а затем привязываем их в редакторе. Мы можем находить булево значение, определяющее сомкнуты ли пальцы (указательный с большим) и силу сжатия пальцев, сохраняя ее в переменную типа float, как показано на рисунке 3.

```
public OVRHand l;
public OVRHand r;

isP = l.GetFingerIsPinching(OVRHand.HandFinger.Index);
lSt = l.GetFingerPinchStrength(OVRHand.HandFinger.Index);
```

Рисунок 3. Функции инициализации рук и работы с ними

Для телепортации мы можем сохранить положение кончика указательного пальца левой руки и телепортировать игровой аватар по направлению нормали к этому кон-

чику (если сомкнуты пальцы). Этого можно добиться с помощью использования функций, как показано на рисунке 4.

```
tipTr = GameObject.Find("Hand_IndexTip").transform;
avatar.transform.Translate(-indexTipL.right * distance);
```

Рисунок 4. Передвижение игрового аватара

Реализованный мультимодальный интерфейс можно представить в виде следующей

схемы (рисунок 5).

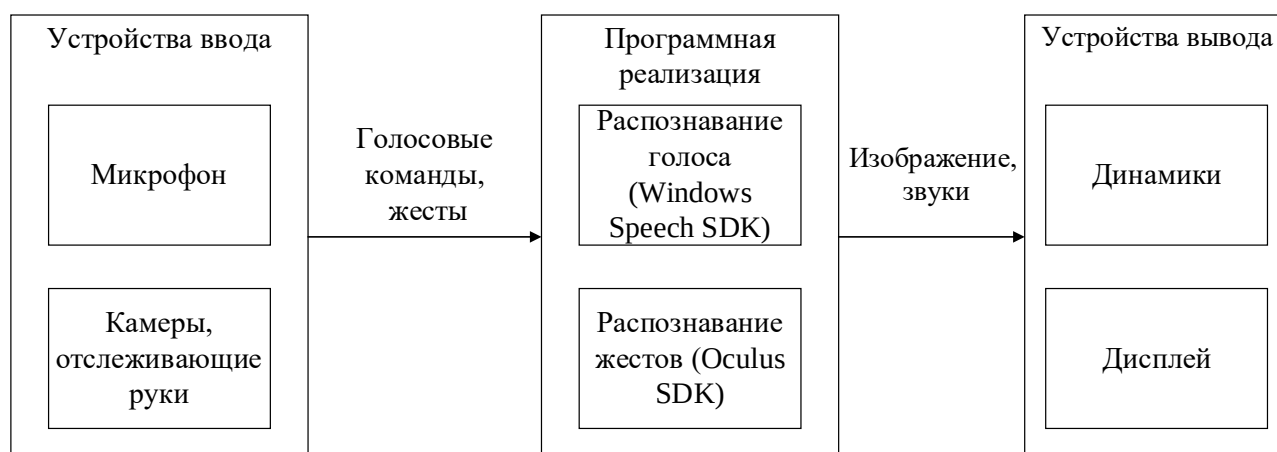


Рисунок 5. Схема мультимодального интерфейса

В таблице 1 представлена возможная реализация синергии компонентов мультимо-

дального интерфейса.

Таблица 1

СИНЕРГИЯ ЖЕСТОВ И ГОЛОСОВЫХ КОМАНД

Жест	Голосовая команда	Результат
Указание пальцем	Телепортация	Телепортация в направлении пальца
Открытая ладонь	Заклинание	Создание магии из ладони, в направлении нормали
Сомкнутые пальцы	Удалить	Выбор объекта и его удаление

Использование предложенных средств ввода имеет широкое применение в VR продуктах. Цель всех VR приложений – создание иммерсивной среды, заменяющей собой привычные информационные системы, увеличение погружения пользователя в виртуальную реальность. Повышение иммерсивности расширяет область применения и повышает качество пользовательского опыта.

Рассмотрим возможность использования средств ввода в игровой системе в виртуальной реальности, где можно использовать заклинания и совершать телепортацию. Среди стандартных инструментов могли бы использоваться пользовательские интерфейсы, в которых можно было бы выбрать нужное заклинание и затем активировать его кнопкой на контроллере. В представленном решении предусмотрено отсутствие пользовательского интерфейса, что является более реалистичным, – достаточно прочесть заклинание. Также в стандартных реализациях VR часто встречается телепортация, осуществляемая с использованием кнопки контроллера, или ходьба – с использованием аналогового стика. Чтобы не нагружать пользователя дополнительными физическими средствами ввода, избавление от контроллеров положительно повлияет на иммерсивность, что и было реализовано через отслеживание рук и жестов [6].

Вследствие того, что VR продукты применяются во многих сферах, улучшение пользовательского опыта через повышение иммерсивности положительно скажется как на продуктах, так и на сферах применения [4]. Более того, большинство индустрий, специализирующихся на виртуальной реальности, являются игровыми, поэтому всё больше людей интере-

суется VR играми, где погружение является одним из главных приоритетов [9].

В настоящей работе предложено и продемонстрировано внедрение мультимодального интерфейса, которое предоставляет пользователям реалистичный опыт в интуитивно понятной структуре при взаимодействии с виртуальной реальностью (с использованием жестов рук, голосового ввода). Предлагаемые способы ввода расширяют возможности разработчиков при создании программного обеспечения для виртуальной реальности и позволяют пользователям погрузиться в виртуальную реальность с новой силой. Например, можно почувствовать себя волшебником, который произносит заклинание и создает его прямо из своей руки. Для управления в виртуальной реальности больше не нужны контроллеры, ведь можно использовать жесты и голос. Таким образом, исчезает физическое средство ввода, встающее преградой к иммерсивности.

Конечно, у предложенных решений есть ограничения, связанные с возможностями аппаратуры, программного обеспечения и достижения реалистичности. Во-первых, не у каждого шлема виртуальной реальности есть микрофон, и у куда меньшего количества моделей есть встроенные камеры для поддержки отслеживания рук. В данной работе был использован VR-шлем Oculus Quest 2, на котором есть 4 камеры для отслеживания, а также Oculus SDK, который позволяет использовать функции и объекты, чтобы упростить работу с отслеживанием рук, который был использован при создании проекта в игровом движке Unity. Во-вторых, на данный момент не существует общепринятого реше-

ния для иммерсивного перемещения игрового аватара, и, хотя наше решение, заключающееся в телепортации аватара с помощью жестов руки, является более иммерсивным, по сравнению с использованием контроллеров, оно, строго говоря, нереалистично, ведь в реальной жизни люди передвигаются с помощью ног или транспортного средства. Отметим также, что для запуска решения нужен компьютер, минимальные требования к которому следующие: четырёхъядерный процессор, 8 гигабайт оперативной памяти, операционная система Windows 10 или выше, видеокарта уровня NVIDIA GeForce GTX 970, ее эквивалент или лучше, Pixel Shader и

Vertex Shader версии не ниже 5.1, видеопамять – 3 гигабайта. Кроме того, должно быть установлено и подготовлено само программное обеспечение вместе с драйверами, обеспечивающее активную связь компьютера с шлемом и функции отслеживания рук (например, для этого может быть использован Oculus Link).

Прямым развитием данной реализации мультимодального интерфейса является нейрокомпьютерный интерфейс, который устанавливает контакт между человеческим мозгом и компьютером, заменяющим обычные пути вывода сигналов в мозг – периферийную нервную систему и мышечную ткань [3].

СПИСОК ЛИТЕРАТУРЫ

1. *Peng C., Hansberger J.T., Cao L., Shanthakumar V.A.* Hand gesture controls for image categorization in immersive virtual environments // IEEE Annual International Symposium Virtual Reality. – Los Angeles: IEEE, 2017. – PP. 331-332.
2. *Zimmerer C., Fischbach M., Latoschik M.E.* Space Tentacles – Integrating Multimodal Input into a VR Adventure Game // IEEE Annual International Symposium Virtual Reality. – Tuebingen: IEEE, 2018. – PP. 745-746.
3. *Wen D., Liang B., Zhou Y., Chen H., Jung T.-P.* The Current Research of Combining Multimodal Brain-Computer Interfaces With Virtual Reality // IEEE Journal of Biomedical and Health Informatics. – 2021. – Vol. 25. – № 9. – PP. 3278-3287.
4. *Lamberti F., Manuri F., Paravati G., Piumatti G., Sanna A.* Using Semantics to Automatically Generate Speech Interfaces for Wearable Virtual and Augmented Reality Applications // IEEE Transactions on Human-Machine Systems. – 2017. – Vol. 47. – № 1. – PP. 152-164.
5. *Jerald J.* Human-centered design for immersive interactions // IEEE Annual International Symposium Virtual Reality. – Los Angeles: IEEE, 2017. – PP. 431-432.
6. *Soares L.P. et al.* Title of tutorial: Designing immersive VR systems: From bits to bolts // IEEE Annual International Symposium Virtual Reality. – Costa Mesa: IEEE, 2012. – PP. 1-6.
7. *Betz M., Hild H.* Language models for a spelled letter recognizer // International Conference on Acoustics, Speech, and Signal Processing (ICASSP). – Detroit: IEEE, 1995. – PP. 856-859.
8. *Nooruddin N., Dembani R., Maitlo N.* HGR: Hand-Gesture-Recognition Based Text Input Method for AR/VR Wearable Devices // IEEE International Conference on Systems, Man and Cybernetics. – Toronto: IEEE, 2020. – PP. 744-751.
9. *Lee S., Park K., Lee J., Kim K.* User Study of VR Basic Controller and Data Glove as Hand Gesture Inputs in VR Games // International Symposium on Ubiquitous Virtual Reality, ISUVR. – Nara: IEEE, 2017. – PP. 1-3.
10. *Machidori Y., Takayama K., Sugita K.* Implementation of multi-modal interface for VR application // International Conference on Awareness Science and Technology (iCAST). – Morioka: IEEE, 2019. – PP. 1-4.

DEVELOPMENT OF AN IMMERSIVE GAME SPACE IN VIRTUAL REALITY USING A MULTIMODAL INTERFACE

SENILOV Mikhail Andreevich

Grand PhD in Technical Sciences, Professor

GAZIZULIN Marat Ruslanovich

student

M.T. Kalashnikov's Izhevsk State Technical University
Izhevsk, Russia

A new approach to the creation of immersive gaming virtual spaces is considered, which consists in the development of a modern, immersive multi-modal interface that uses the principle of voice input and hand gesture processing. When creating an immersive space, advanced technologies are used, which condition the synergy of the elements of a multimodal interface, providing a new level of quality for virtual reality software.

Key words: virtual reality, immersiveness, multimodal interface, hand gesture processing, voice recognition, computer game development.
